REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 27-01-2012	**2. REPORT TYPE** Final	**3. DATES COVERED** *(From - To)* 1 May 2004 - 30 Apr 2010

4. TITLE AND SUBTITLE
Chaotic Models and Anomaly Detection for Complex Data Networks

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-04-1-0319

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Brian R. Hunt, Edward Ott, James A. Yorke

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Maryland
ORAA
3112 Lee Building
College Park MD 20742

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
USAF, AFRL
AFOSR
4015 Wilson Blvd., Room 713
Arlington VA 22203-1954

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR/PK3

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2012-0112

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Our main goal was to detect and describe deterministic aspects of traffic behavior in data networks, in order to provide a basis for better detection of anomalous network activity. We also sought to characterize the robustness of complex data networks to (possibly malicious) perturbations, in order to help engineer against disruptions.

Throughout this project we have developed dynamical systems models for TCP network traffic on networks of increasing complexity, guided by real packet-level data and network simulation software. We also developed techniques for estimating the network state (e.g., router queue sizes and round-trip times of data flows) from packet-level data. We investigated methods for short-term prediction of ``normal'' network activity to use as a baseline for anomaly detection. We modeled peer-to-peer network activity and developed methods for detecting such activity. Finally, we examined the stability of TCP network dynamics and their response to perturbations that could be used as low-volume denial-of-service attacks. We found large-scale network dynamics to be robust to such perturbations, but identified mechanisms for localized disruptions.

15. SUBJECT TERMS
TCP network modeling, network state characterization

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Brian R. Hunt
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU		**19b. TELEPHONE NUMBER** *(include area code)* 301-405-5056

Final Report for Grant #FA9550-04-1-0319

Brian Hunt (PI), Edward Ott, and James Yorke
University of Maryland, College Park MD 20742

Objectives

Our main goal was to detect and describe deterministic aspects of traffic behavior in data networks, in order to provide a basis for better detection of anomalous network activity. We also sought to characterize the robustness of complex data networks to (possibly malicious) perturbations, in order to help engineer against disruptions.

Methods/Findings

We developed and published [2,3] a method for estimating, from packet data collected at a single point in a real network, quantities such as the current round trip time and TCP congestion window for individual sender-receiver pairs. We found that we can use our measurements to predict the timing of packet drops; for example, predictions based on our estimates of round trip time are significantly better than predictions based only on the statistics of inter-drop times. In conjunction with our models, discussed below, this work enables short-term prediction of events like congestion in a particular part of the network.

Precise modeling of packet-level network dynamics requires several variables to track even a single TCP flow through a router. We found that we can capture the dynamics of the model reasonably well using a single variable for each flow, the size of its congestion window. In particular, the same dynamical phenomena that occur in packet-level simulations occur in our simplified model. By simplifying the model, we gained a better understanding of why these phenomena occur, and we were able to scale the model to much larger networks.

We published [1] a model for TCP traffic (with RED congestion control) describes the dynamics of a network in terms of two main state variables: the size of the congestion window for each data flow between a given sender-receiver pair, and the size of the queue at each router. In addition, it keeps track of a filtered (time-averaged) queue size for each router to simulate the RED mechanism for dropping packets, and has some timer variables to track the effect of packet loss. We found that even in a network consisting of a single sender, router,

and receiver, this model can behave chaotically (unpredictable in the long run despite being deterministic) in response to a periodic input to the router from a non-TCP source.

We also published [6] a study of bifurcations in our model, and in the network simulator `ns`, for small networks. In this work, we investigated the response of a small network to periodic short packet bursts. Our motivation was to understand the possible effects of a low-volume denial of service attack, where a host tries to disrupt a network with inobtrusive but well-timed packet bursts rather than just flooding the network with packets. We found that as the period of the bursts varies there are sudden changes in the network response between stable periodic behavior and chaotic behavior, and that these bifurcations are qualitatively reproduced by a one-dimensional model that tracks the phase of congestion events relative to the periodic bursts.

To study large networks, we developed and published [5] a streamlined model of TCP-RED network traffic that exhibits similar dynamics to our previous model, but requires only keeping track of a rate (proportional to the congestion window) for each data flow. This gave rise to a piecewise linear model, with complexity arising from the variety of sequences in which different routers on the network can become congested and drop packets. We found that regardless of network size, the dynamics are stable and periodic for typical parameter values. The model assumes that the network traffic is dominated by bulk flows, in which a sender is transmitting information to a receiver as fast as the TCP protocol allows for an extended period of time. The parameters are the round-trip times of these flows and the throughput capacities of the routers involved, and the initial conditions are the send rates of the different flows at a particular time. Our results suggest a large-scale robustness of TCP networks to small perturbations, despite the potential for local disruptions indicated by our lower-level model.

Finally, we investigated [4] routing and load-balancing algorithms for peer-to-peer networks, in order to understand and help detect the dynamics of file-sharing traffic. We also examined sets of network trace data, in which much of the peer-to-peer activity is easy to identify due to the port numbers used. We used this data to test methods for classifying network traffic without considering port numbers (since covert peer-to-peer activity will use ports normally associated with other types of traffic). We found it most fruitful to identify quantities that should correlate to peer-to-peer activity based on our understanding of its dynamics, then use standard algorithms for decision tree analysis to determine how to classify network traffic according to these quantities. An example of such a quantity is the frequency with which a particular computer receives data from one computer and then soon after initiates a connection with a different computer.

Personnel supported

Faculty: Brian Hunt, Edward Ott, James Yorke
Graduate Students: Ian Frommer, Ryan Lance, Amy Finkbiner, Russell Halper, Yiwei Chen

Publications

1. I. Frommer, R. Lance, B. R. Hunt, E. Ott, J. A. Yorke, E. Harder, "Modeling congested internet connections", Proceedings of the Second IASTED International Conference on Computer and Communications Networks (November 2004, Cambridge MA), pp. 319–324.

2. R. Lance, I. Frommer, B. R. Hunt, E. Ott, J. A. Yorke, E. Harder, "Round-trip time inference via passive monitoring", Proceedings of the Workshop on Large Scale Network Inference (LSNI): Methods, Validation, and Applications, ACM SIGMETRICS (June 2005, Banff, Alberta, Canada).

3. R. Lance, Network State Estimation Via Passive Traffic Monitoring, Ph.D. thesis (2005), `http://drum.lib.umd.edu/handle/1903/2429`.

4. A. Finkbiner, Global Phenomena from Local Rules: Peer-to-Peer Networks and Crystal Steps, Ph.D. thesis (2007), `http://drum.lib.umd.edu/handle/1903/7682`.

5. R. D. Halper, E. J. Harder, B. R. Hunt, J. A. Yorke, Stability of TCP dynamics in large data networks, SIAM J. Appl. Dynam. Sys. **8** (2009), 146–159.

6. I. Frommer, E. Harder, B. Hunt, R. Lance, E. Ott, J. Yorke, Bifurcation and chaos in a periodically probed computer network, Int. J. Bif. Chaos **19** (2009), 3129–3141.

Interactions/Transitions

Most of our work on this project has been done in collaboration with Eric Harder of NSA. Ryan Lance received his Ph.D. in 2005 and is now working at NSA. Ian Frommer received his Ph.D. in 2005 and is now teaching at the Coast Guard Academy. Amy Finkbiner received her Ph.D. in 2007 and is now working for Lockheed Martin. Russell Halper finished his Ph.D. in 2009 and is now working for Nestlé. Yiwei Chen has transfered to Princeton University.